HHS Public Access

Author manuscript

Radiology. Author manuscript; available in PMC 2017 November 27.

Published in final edited form as:

Radiology. 2017 September ; 284(3): 777–787. doi:10.1148/radiol.2017161736.

Effect of Radiation Dose Reduction and Reconstruction Algorithm on Image Noise, Contrast, Resolution, and Detectability of Subtle Hypoattenuating Liver Lesions at Multidetector CT:

Filtered Back Projection versus a Commercial Model–based Iterative Reconstruction Algorithm[1]

Justin Solomon, PhD, **Daniele Marin, MD**, **Kingshuk Roy Choudhury, PhD**, **Bhavik Patel, MD**, and **Ehsan Samei, PhD**

[1]Carl E. Ravin Advanced Imaging Laboratories, Department of Radiology, Duke University Medical Center, 2424 Erwin Rd, Suite 302, Durham, NC 27705

Abstract

Purpose—To determine the effect of radiation dose and iterative reconstruction (IR) on noise, contrast, resolution, and observer-based detectability of subtle hypoattenuating liver lesions and to estimate the dose reduction potential of the IR algorithm in question.

Materials and Methods—This prospective, single-center, HIPAA-compliant study was approved by the institutional review board. A dual-source computed tomography (CT) system was used to reconstruct CT projection data from 21 patients into six radiation dose levels (12.5%, 25%, 37.5%, 50%, 75%, and 100%) on the basis of two CT acquisitions. A series of virtual liver lesions (five per patient, 105 total, lesion-to-liver prereconstruction contrast of −15 HU, 12-mm diameter) were inserted into the raw CT projection data and images were reconstructed with filtered back projection (FBP) (B31f kernel) and sinogram-affirmed IR (SAFIRE) (I31f-5 kernel). Image noise (pixel standard deviation), lesion contrast (after reconstruction), lesion boundary sharpness (average normalized gradient at lesion boundary), and contrast-to-noise ratio (CNR) were

Address correspondence to J.S. (justin.solomon@duke.edu).

Online supplemental material is available for this article.

Author contributions:

Guarantors of integrity of entire study, J.S., E.S.; study concepts/study design or data acquisition or data analysis/ interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; manuscript final version approval, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, J.S., B.P., E.S.; clinical studies, D.M., B.P.; experimental studies, J.S., E.S.; statistical analysis, J.S., K.R.C.; and manuscript editing, J.S., D.M., B.P., E.S.

Conflicts of interest are listed at the end of this article.

Disclosures of Conflicts of Interest: J.S. Activities related to the present article: received a grant from Siemens Medical Solutions. Activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships. **D.M.** Activities related to the present article: received research support from Siemens Medical Solutions. Activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships. **K.R.C.** disclosed no relevant relationships. **B.P.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: received personal fees from GE Healthcare. Other relationships: disclosed no relevant relationships. **E.S.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: received a grant from Siemens Medical Systems and GE Healthcare. Other relationships: disclosed no relevant relationships.

compared. Next, a two-alternative forced choice perception experiment was performed (16 readers [six radiologists, 10 medical physicists]). A linear mixed-effects statistical model was used to compare detection accuracy between FBP and SAFIRE and to estimate the radiation dose reduction potential of SAFIRE.

Results—Compared with FBP, SAFIRE reduced noise by a mean of 53% ± 5, lesion contrast by 12% ± 4, and lesion sharpness by 13% ± 10 but increased CNR by 89% ± 19. Detection accuracy was 2% higher on average with SAFIRE than with FBP ($P = .03$), which translated into an estimated radiation dose reduction potential (±95% confidence interval) of 16% ± 13.

Conclusion—SAFIRE increases detectability at a given radiation dose (approximately 2% increase in detection accuracy) and allows for imaging at reduced radiation dose (16% ± 13), while maintaining low-contrast detectability of subtle hypoattenuating focal liver lesions. This estimated dose reduction is somewhat smaller than that suggested by past studies.

Iterative reconstruction (IR) algorithms in multidetector computed tomography (CT) have been shown to improve image quality compared with the traditional filtered back projection (FBP) algorithms (1–12) and may allow for imaging at reduced radiation doses without compromising image quality. The sinogram-affirmed IR (SAFIRE) algorithm (Siemens Medical Solutions, Forchheim, Germany) has been available for several years and studied extensively in the context of phantoms (1,4,6,12–15), simulations (16,17), and clinical data (7,9,18–26). These studies have concluded that SAFIRE has significant (25%–75%) radiation dose savings potential but were limited in one or more of the following ways.

First, by ethical necessity most clinical studies have considered only two radiation dose levels (21–26), making it difficult to understand the full functional relationship between radiation dose and image quality (for a specific task) based on only two data points. For example, a common study design is a noninferiority-based comparison between FBP at full radiation dose and IR at some arbitrarily chosen reduced radiation dose. With such data, it is difficult to estimate the potential radiation dose reduction from the iterative algorithm with any amount of precision, and it is impossible to estimate the performance of the algorithm for radiation dose levels that were not explicitly tested.

Second, many studies have not used objective measures of image performance (eg, sensitivity, specificity, detection accuracy) as figures of merit. For example, the clinical studies have typically been based on subjective visual scoring of quality (7,21,22,24,26) and the phantom studies have been based on the measurement of physical image characteristics such as noise and/or resolution (1,6,13,15). Under the formalism established by International Commission on Radiation Units and Measurements report 54 (27), image quality is defined as the effectiveness by which an image can be used for its intended task. Thus, when comparing two reconstruction algorithms, it is important to base this comparison on how well images from each reconstruction can objectively be used for a clinically relevant task (eg, detection of subtle lesions). This is especially true when comparing algorithms that may have varying noise texture, contrast, and resolution properties because detectability is affected by all these factors simultaneously (4,28). In other words, what is most important is not how much noise is reduced by an IR algorithm but how much the detectability is increased.

Third, the phantom and simulation studies have relied on data that are fundamentally oversimplified compared with patient data. For example, Xu et al (16) performed an excellent simulation study in which raw CT projection data of virtual anthropomorphic phantoms (with subtle lesions) were simulated and images were reconstructed by using FBP and SAFIRE algorithms from Siemens. Objective detectability was assessed by using a mathematical observer model. The major limitation of such a study is that the virtual phantoms used were essentially uniform within a given organ. This leads to simulated data that do not contain the level of heterogeneity and complexity encountered in clinical work. It has been shown that owing to nonlinear regularization techniques, IR algorithms behave differently in uniform regions compared with regions with variable anatomic features and textures (11,29–32). Thus, it is questionable whether the results from studies based on uniform phantoms can translate to how an IR algorithm would perform with real clinical data.

The purpose of this study was to determine the effect of radiation dose and reconstruction algorithm (FBP vs SAFIRE) on noise, contrast, and observer-based detectability of subtle hypoattenuating liver lesions. The aforementioned limitations of previous studies (too few dose levels, poor metrics of image quality, and/or over-simplified phantom and/or simulation data) were overcome by using so-called hybrid images (ie, real CT images infused with virtual lesions) along with a dose-splitting technique that enables imaging of the same patient at up to six radiation dose levels with just two CT acquisitions. With this paradigm, we aimed to objectively and directly estimate how radiation dose and reconstruction algorithm affect detectability, thus enabling a realistic estimate of the radiation dose reduction potential of SAFIRE for the given clinical task.

Materials and Methods

This study was financially supported by Siemens Medical Solutions USA (Malvern, Pa). The authors had complete control of the data and information submitted for publication. This prospective, single-center, Health Insurance Portability and Accountability Act–compliant study was approved by the institutional review board of Duke University.

Image Acquisition

Twenty-one patients (mean age ± standard deviation, 60 years ± 11; median age, 62 years; age range, 44–80 years), including 10 men (mean age, 66 years ± 10; median age, 64 years; age range, 49–80 years) and 11 women (mean age, 56 years ± 9; median age, 53 years; age range, 44–69 years), with clinical indication for metastatic colorectal cancer were prospectively enrolled on the basis of known or suspected liver metastases as indicated at previous multidetector CT ($n = 15$) or ultrasonographic ($n = 5$) examinations or increased carcino-embryonic antigen tumor marker levels (>5 ng/mL) ($n = 6$). Exclusion criteria included a clinical history of intolerance to iodinated contrast material or renal failure defined as serum creatinine level greater than 2.0 mg/mL (177 μmol/L). No patient met the exclusion criteria in our study. The mean patient effective diameter—measured digitally on the midabdomen of the scout radiograph and calculated according to American Association of Physicists in Medicine report 204 (33)—was 27.6 cm ± 4.7 (median, 27.8 cm; range,

20.3–39.1 cm). Note that this range of patient sizes does not include the largest patients typically encountered clinically.

Each patient underwent two successive CT acquisitions (scan 1, full dose; and scan 2, half dose) by using modified versions of our institution's standard contrast material–enhanced abdominal protocol with a dual-source CT scanner (Somatom Flash, Siemens Medical Solutions) (Table 1). On this dual-source scanner, it is possible to set each x-ray tube (tube A and tube B) to have the same tube potentials (in kilovolts) but different tube currents (in milliamperes), allowing for simultaneous acquisition of data from the same patient at multiple radiation dose levels. It is further possible to reconstruct images on the basis of either data from each individual tube or combined data from both tubes. Thus, a single CT acquisition results in CT images at three radiation dose levels (tube A, tube B, and tubes A and B). With two CT acquisitions, a total of six radiation dose levels were achieved per patient. Each of the two CT acquisitions split the baseline tube current 75% to tube A and 25% to tube B, resulting in images representing 100% (scan 1, tubes A and B), 75% (scan 1, tube A), 50% (scan 2, tubes A and B), 37.5% (scan 2, tube A), 25% (scan 1, tube B), and 12.5% (scan 2, tube B) radiation dose relative to our standard clinical radiation dose for an abdominal protocol. The scan range in scan 2 was limited to include only the liver, thus minimizing the overall radiation burden of scan 2 in terms of the dose-length product. It was estimated that this second scan increased the total radiation burden to the patient by about 15% relative to a standard of care diagnostic examination when accounting for scan coverage. In terms of volume CT dose index and size-specific dose estimate, the 100% dose level corresponded to a mean of 10.3 mGy ± 3.1 and 13.0 mGy ± 2.5, respectively (variability due to the use of tube current modulation).

All scans were obtained in a single breath hold during the hepatic parenchymal (portal venous dominant) phase (70 seconds after the start of contrast material injection). Scan 1 was obtained in the craniocaudal direction, whereas scan 2 was obtained immediately after in the opposite direction to minimize contrast material timing differences between the two acquisitions. The bolus of contrast material consisted of 150 mL with an iodine concentration of 300 mg iodine per milliliter (iopamidol [Isovue 300; Bracco Diagnostics, Princeton, NJ]) and was administered with an 18–20-gauge cannula inserted into an antecubital fossa vein by using a dual-chamber mechanical power injector (Empower; E-Z-Em, Lake Success, NY) at a flow rate of 3 mL/sec.

The raw projection data were exported from the scanner console onto an external hard drive. After inserting virtual lesions into the projection data (described in the next section), images were reconstructed with FBP (B31f kernel) and SAFIRE-5 (I31f kernel) by using an offline reconstruction system (ReconCT, Siemens Medical Solutions). Note that the vendor's implementation of FBP (so-called weighted FBP) has already been shown to be superior to the classic FBP algorithm (34,35). This implementation will be referred to as FBP throughout. Twelve CT datasets were reconstructed per patient (six radiation dose levels × two reconstruction algorithms).

Virtual Lesion Insertion

In practice it is extremely difficult to find clinical cases of subtle lesions that are near the threshold of detectability. However, it is exactly these cases that one would expect to be most sensitive to changing radiation dose and reconstruction algorithm. For this reason, in this study we used a method of inserting subtle virtual lesions into the real CT images, resulting in so-called hybrid images. This made it possible to investigate the effect of radiation dose and reconstruction algorithm on the detection of very subtle lesions in a controlled and targeted manner.

On the basis of a previously published method for generating realistic virtual liver lesions (36), a custom software tool (LesionTool; Duke University, Durham, NC) (Fig 1) was used to generate a series of liver lesion models having fixed inherent (ie, prereconstruction) contrast of −15 HU with variable size and shape (average radius, 6 mm ± 1). Note that although typical clinical liver lesions tend to have higher contrast compared with those used in this study—average of approximately −35 HU on the basis of limited available data (37)—this contrast level was chosen to target lesions that were near the threshold of detectability and to assess the performance of the IR algorithm under the most challenging conditions. Several phantom-based studies have likewise targeted lesions with contrasts in the range of −12 to −40 HU (38,39). Five lesions were generated per patient, resulting in 105 total lesions. Suitable insertion locations were identified as small regions within the liver (just big enough to fit the lesion) but void of obvious features (eg, blood vessels or real lesions). Some lesions were inserted in close proximity to blood vessels or other normal anatomic features; however, they were not superimposed on top of those features. Virtual lesions were not inserted near real lesions. The lesion models were voxelized to match the targeted CT resolution and exported as solitary three-dimensional image volumes. These image volumes were fed into a projection-based lesion insertion routine (36,40) developed as a plug-in to the ReconCT software package (Siemens Medical Solutions). This routine used ReconCT's underlying forward projector to create synthetic projection images of the lesions. These synthetic projections were added (ie, superimposed) onto the real patient projection data. Because the underlying geometry of the CT system is known exactly by ReconCT, this method enabled the synthesis of realistic projection data without artifacts. The altered projection data were then reconstructed by using the manufacturer's standard reconstruction methods. Because the lesions were inserted before reconstruction, any nonlinear effects from the reconstruction algorithm were faithfully represented in how the final lesion was rendered in the reconstructed image space. Similar projection-based lesion insertion techniques have been shown to produce realistic results (41). Corresponding images with and without lesions were reconstructed with both FBP and SAFIRE. From these image volumes, 40-mm regions of interest (ROIs) were extracted around the lesion locations from both the lesion-present and lesion-absent images.

Noise, Contrast, and Resolution

A series of physical measurements were made on the image ROIs, including noise magnitude, contrast, contrast-to-noise ratio (CNR), and resolution (ie, lesion sharpness). For the measurements below, sub-ROIs were defined having the shape of the inserted lesion and being contained within the inserted lesion as illustrated in Figure 2. Noise magnitude was

taken as the standard deviation of pixel values within this sub-ROI region from the lesion-absent image. Contrast was estimated by first computing a difference image, I_{diff}, between the lesion-present and lesion-absent images. The contrast was taken as the mean pixel value of I_{diff} within the aforementioned sub-ROI region. CNR was computed as the ratio between the noise magnitude and the contrast. Finally, the resolution was assessed by computing the average relative gradient magnitude, $|\nabla|$, for pixels in I_{diff} near the boundary of the lesion (see Fig 2). $|\nabla|$ is defined as the change in pixel value (normalized by the maximum pixel value in the image) per millimeter. A high gradient implies a sharp transition between the lesion and the background, and a low gradient implies a blurry boundary.

Detectability

In addition to the signal-present ROIs used earlier, a series of 20 lesion-absent ROIs were identified in different locations (compared with the lesion insertion locations) for each patient. On the basis of this ensemble of lesion-present and lesion-absent ROIs, a two-alternative forced choice (2AFC) detection experiment was performed to assess the effect of radiation dose and reconstruction algorithm on detection accuracy (Fig 3). Sixteen readers (six radiologists, 10 medical physicists) participated. For each radiation dose and reconstruction condition, readers were shown 105 trials (ie, image pairs) in randomized order for a total of 1260 2AFC trials per reader. For each trial, the reader was shown two images, one with the lesion and one without. Their task was to select the image most likely to contain the lesion. The reader outcomes (detection or nondetection) were recorded for each trial. In addition, to minimize any memory effect of seeing the same lesion as rendered in different radiation dose and/or reconstruction conditions, images were flipped and rotated randomly before being shown to the observers.

Statistical Analysis

Noise, contrast, and CNR were compared pairwise between FBP and SAFIRE by using linear regression analysis. The slope and corresponding 95% confidence interval (CI) of the regression line were used to assess the effect of SAFIRE on a given quantity (slope > 1 implies the quantity is greater in SAFIRE images, slope = 1 implies the quantity is similar, and slope < 1 implies it is greater in FBP images). The coefficient of determination (R^2) was used as a goodness-of-fit metric.

The binary 2AFC reader outcomes (ie, detection or nondetection) were statistically analyzed by fitting a generalized linear mixed effects regression model (binomial distribution with a probit link function). The probability of detection (ie, detection accuracy) was modeled as follows:

$$P\left(Y_{i,j,k,n}=1\right)=\Phi\left(\mu+\alpha\ln\left(D_k\right)+\beta S_i+R_j+L_n+\varepsilon\right),$$

where $Y_{i,j,k,n}$ is the reader outcome (1 for detection, 0 for nondetection) for the ith reconstruction algorithm (i = 1 corresponds to FBP, i = 2 corresponds to SAFIRE), jth reader (j = 1,…,16), kth dose level (k = 1,…,6), and nth lesion (n = 1,…,105). Here $\Phi(x)$ is the standard normal cumulative distribution function, μ a baseline intercept term (representing FBP at 1% radiation dose), α a radiation dose term, $\ln(D_k)$ the natural logarithm of the kth

radiation dose level, β a reconstruction algorithm term, S_i a binary dummy variable encoding the reconstruction algorithm (FBP corresponds to $S_1 = 0$, SAFIRE corresponds to $S_2 = 1$), R_j a random effect to account for reader variability, L_n a random effect to account for variable inherent conspicuity across the ensemble of lesions, and ε an error term. All random terms were assumed to have zero mean normal distributions with standard deviations of σ_R, σ_L, and σ_ε, respectively. Model terms with $P < .05$ were considered indicative of a statistically significant difference. This model was chosen due to the theoretical relationship between radiation dose and an observer's detectability index (exponentially related) and the relationship between detectability index and detection accuracy (related by $\Phi[x]$) (10,42). The log of radiation dose was taken to linearize the data with respect to radiation dose, thus better satisfying the model assumptions. This analysis used the Statistics and Machine Learning Tool-box in MATLAB (*fitglme ()* function).

This model was used to compare the mean detection accuracy (averaged across all readers and lesions) between FBP and SAFIRE and to estimate the radiation dose reduction potential of SAFIRE, defined as the reduced radiation dose at which SAFIRE had equivalent average detection accuracy compared with FBP at 100% reference radiation dose. Given the statistical model parameters, the relative radiation dose reduction potential of SAFIRE, ΔD, was analytically computed as follows:

$$\Delta D \equiv 100\% \cdot \frac{D_F - D_S}{D_F} = 100\% \cdot 1 - e^{\frac{-\beta}{\alpha}},$$

where D_F is the reference FBP radiation dose and D_S the reduced radiation dose at which SAFIRE had equivalent detection accuracy. This equation can be derived directly from the statistical model by setting $\bar{A}_F(D_F) = \bar{A}_S(D_S)$, solving for D_S, and then substituting D_S into the definition of relative radiation dose reduction given above. Here $\bar{A}_F$ and $\bar{A}_S$ are the modeled detection accuracy from FBP and SAFIRE, respectively, averaged across readers and lesions. The standard error of ΔD, $\sigma_{\Delta D}$, was computed by propagating the standard errors of α and β (as output from the statistical model) through the above equation (ignoring high-order terms) and 95% CIs of ΔD were computed as $\Delta D \pm 1.96\sigma_{\Delta D}$.

Results

A montage of the ROI images is shown in Figure 4. On average (across all radiation dose levels), SAFIRE reduced the noise by 53% ± 5, reduced contrast by 12% ± 4 (1–2 HU), increased CNR by 89% ± 19, and reduced lesion sharpness by 13% ± 10 compared with FBP. The linear regression lines of SAFIRE versus FBP had slopes of 0.46 (95% CI: 0.45, 0.47), 0.88 (95% CI: 0.86, 0.89), 1.82 (95% CI: 1.78, 1.86), and 0.64 (95% CI: 0.62, 0.66) for noise, contrast, CNR, and lesion sharpness, respectively. R^2 was 0.90, 0.95, 0.93, and 0.86, respectively (Fig 5).

Results from the 2AFC detection experiments are given in Table E1 (online), and the output of the statistical model is given in Table 2. Detection accuracy increased with increasing dose ($P < .001$) and with use of SAFIRE (2% increase on average, $P = .03$). The effect of radiation dose on detection accuracy was relatively mild at high radiation doses (2%

decrease in detection accuracy between 100% and 75% radiation dose) and more drastic at lower radiation doses (5% decrease in detection accuracy between 25% and 12.5% radiation dose). There was also considerable variability in the data owing to readers and lesions, with standard deviations of these effects being comparable to the magnitude of the fixed effects in the statistical model (Table 2). Compared with the reference 100% clinical radiation dose with FBP, SAFIRE had equivalent performance at 84% radiation dose, which implies a dose reduction potential (ΔD) (±95% CI) of 16% ± 13, as illustrated in Figure 6.

Discussion

Our data provide further evidence that SAFIRE could facilitate imaging of subtle hypoattenuating focal liver lesions at lower radiation doses. However, the magnitude of this radiation dose reduction potential (approximately 16%) is smaller than that suggested by previous investigators (1,9,16) and is mitigated by reduced contrast and resolution for low-contrast detection tasks. To our knowledge, this is the first study to explicitly measure detectability of subtle lesions on patient images in an objective fashion by using human readers over a wide range of radiation dose levels. Previous studies have relied on phantom-based measures, mathematical observer models, or subjective scoring of image quality. A major advantage of the methodology used in this study is that the functional relationship between radiation dose and detection accuracy of human readers was ascertained, allowing for a granular estimation of the magnitude of radiation dose reduction potential from SAFIRE along with an understanding of how sensitive the detection task was to changing radiation dose. Note that in this study, as with any human reader-based study, there was considerable interreader variability. There was also considerable interlesion variability (ie, some lesions were more difficult to detect than others), as evidenced by the random terms of the statistical model. The analysis of the data was primarily focused on determining the radiation dose reduction potential of SAFIRE for the “average” reader, further averaged across an ensemble of lesions. Due to the aforementioned interreader variability, it is possible that individual readers would respond differently to changing radiation dose and reconstruction algorithm conditions. Future work could explore this variability further.

The data show that reducing the radiation dose (with everything else held constant) will have a negative effect on detectability. This implies that to achieve a certain level of detection accuracy for subtle lesions, a minimum dose is required. Also of note is that the functional relationship between radiation dose and detection accuracy shows a relatively high slope at low radiation doses (12.5%–37.5%) but a rather shallow slope at high radiation doses (75%–100%), meaning that detection accuracy is most sensitive to changing radiation dose at lower radiation dose levels. This functional relationship follows what would be predicted by the physics and/or statistics of CT imaging along with statistical signal detection theory and is not a novel result of this study but is an important point to re-emphasize for those attempting radiation dose optimization. It was only possible to demonstrate this functional relationship due to the CT scanner’s ability to acquire data from the same patient at different radiation dose levels within a very narrow time window.

The observed effect of SAFIRE on physical image properties is mostly consistent with previous studies showing a sizeable noise reduction from SAFIRE (1,4,6,7,9,12–18,21–

26,43). Of note, however, is the fact that, in this study, we also observed a reduction in lesion contrast and resolution from SAFIRE. Phantom studies from Solomon et al (11) have demonstrated that a similar reconstruction algorithm from the same vendor (Advanced Modeled Iterative Reconstruction–ADMIRE, Siemens Medical Solutions) has highly variable spatial resolution as a function of contrast. For example, ADMIRE had poorer resolution than FBP for very low contrast features but better resolution for high contrast features (11). In general, poorer resolution can result in a loss of contrast. Thus, it is not surprising that SAFIRE seems to exhibit similar behavior. Similar results have also been reported for iterative algorithms from other vendors (44).

This also highlights the strength of using objective task-based image quality metrics when comparing different reconstruction algorithms that potentially have different noise, contrast, and resolution properties. Radiation dose reduction claims on the basis of noise or CNR alone should be received with skepticism. For example, in this study, SAFIRE reduced noise by 52% on average and the data showed that one could image at 87% reduced radiation dose with SAFIRE to achieve similar noise to FBP at full radiation dose. However, on the basis of the 2AFC detection experiment, SAFIRE at 85% reduced radiation dose had approximately 13% lower detection accuracy than FBP at full radiation dose. Similarly, one can image at 75% reduced radiation dose and achieve similar CNR to that with FBP at full radiation dose. However, that would lead to a reduction in the detection accuracy by approximately 8%.

In this study, the virtual lesions were purposefully designed to be near the threshold of detectability to isolate (as much as possible) the effect of the reconstruction algorithm on detectability. Alternatively, the comparison of lesions that are easily detectable on both FBP and IR images provides very little information about the benefits of the IR algorithm. As a result, the lesions in this study were probably subtler than those typically encountered in clinical practice. In addition, in the calculation of the dose reduction potential, it was assumed that the goal was to maintain constant detectability of these subtle lesions. However, it is conceivable that one may be willing to somewhat sacrifice detectability of very subtle lesions to reduce radiation dose for most patients who have more conspicuous lesions or no lesions at all. This should be an active area of discussion within the imaging community moving forward.

A few limitations of our study merit acknowledgment. First, a limited number of patients (n = 21) were used. There were two reasons for this limitation: *(a)* Only a limited number of patients were successfully recruited for the study, and *(b)* there is a practical limit to the number of 2AFC trials that can be shown to readers in a given session before their performance suffers (ie, reader fatigue). In this study, readers completed more than 1200 trials in a single session. Second, the 2AFC paradigm reduces the complex process of image interpretation down to making a binary decision of lesion present or lesion absent, aiming to test the perceived conspicuity of the lesions in small ROIs. The expertise of a radiologist is centered on making a clinical diagnosis given the full image data and other information about the patient. This is a much broader task compared with assessing if a lesion is present in a small ROI image. Thus, the detectability results from this study represent a relatively simple case compared with clinical reality. Nevertheless, these types of experiments do offer valuable information about how easily a subtle lesion might be perceived in a more typical

clinical scenario and how different reconstruction algorithms render the same lesion differently. In addition, the study used a number of imaging physicists who are not explicitly trained in clinical interpretation of CT images. Nonetheless, for the same reasons mentioned earlier (ie, the simplified nature of the detection task at hand), we do not believe this is a significant limitation of the study because imaging physicists are accustomed to performing low-contrast detection tasks as part of routine imaging physics testing. Finally, despite efforts to minimize contrast timing differences, it is still possible that the contrast enhancement of the liver parenchyma was somewhat different between the two consecutive scans. However, because the lesions of interest in this study were inserted virtually, any difference in background liver enhancement would have little to no effect on the results of this study.

In conclusion, the SAFIRE algorithm reduces image noise but also reduces contrast and resolution of subtle hypoattenuating liver lesions compared with FBP. The net effect is that SAFIRE increases detectability by approximately 2% at a given radiation dose and is able to offer equivalent detectability at approximately 16% ± 13 reduced radiation dose. This estimated dose reduction is somewhat smaller than past studies have suggested.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by Siemens USA.

The authors thank Juan Carlos Ramirez-Giraldo, PhD, and Martin Sedlmair, PhD, for their help in reconstructing the image data.

Abbreviations

CI	confidence interval
CNR	contrast-to-noise ratio
FBP	filtered back projection
IR	iterative reconstruction
ROI	region of interest
SAFIRE	sinogram-affirmed IR
2AFC	two-alternative forced choice

References

1. Baker ME, Dong F, Primak A, et al. Contrast-to-noise ratio and low-contrast object resolution on full- and low-dose MDCT: SAFIRE versus filtered back projection in a low-contrast object phantom and in the liver. AJR Am J Roentgenol. 2012; 199(1):8–18. [PubMed: 22733888]
2. Beister M, Kolditz D, Kalender WA. Iterative reconstruction methods in x-ray CT. Phys Med. 2012; 28(2):94–108. [PubMed: 22316498]

3. Chang W, Lee JM, Lee K, et al. Assessment of a model-based, iterative reconstruction algorithm (MBIR) regarding image quality and dose reduction in liver computed tomography. Invest Radiol. 2013; 48(8):598–606. [PubMed: 23511193]

4. Chen B, Ramirez Giraldo JC, Solomon J, Samei E. Evaluating iterative reconstruction performance in computed tomography. Med Phys. 2014; 41(12):121913. [PubMed: 25471973]

5. Gervaise A, Osemont B, Lecocq S, et al. CT image quality improvement using adaptive iterative dose reduction with wide-volume acquisition on 320-detector CT. Eur Radiol. 2012; 22(2):295–301. [PubMed: 21927791]

6. Ghetti C, Palleri F, Serreli G, Ortenzia O, Ruffini L. Physical characterization of a new CT iterative reconstruction method operating in sinogram space. J Appl Clin Med Phys. 2013; 14(4):4347. [PubMed: 23835395]

7. Han BK, Grant KL, Garberich R, Sedlmair M, Lindberg J, Lesser JR. Assessment of an iterative reconstruction algorithm (SAFIRE) on image quality in pediatric cardiac CT datasets. J Cardiovasc Comput Tomogr. 2012; 6(3):200–204. [PubMed: 22682262]

8. Kalender WA, Beister M, Boone JM, Kolditz D, Vollmar SV, Weigel MC. High-resolution spiral CT of the breast at very low dose: concept and feasibility considerations. Eur Radiol. 2012; 22(1):1–8. [PubMed: 21656331]

9. Kalra MK, Woisetschläger M, Dahlström N, et al. Radiation dose reduction with sinogram affirmed iterative reconstruction technique for abdominal computed tomography. J Comput Assist Tomogr. 2012; 36(3):339–346. [PubMed: 22592621]

10. Solomon J, Mileto A, Ramirez-Giraldo JC, Samei E. Diagnostic performance of an advanced modeled iterative reconstruction algorithm for low-contrast detectability with a third-generation dual-source multidetector CT scanner: potential for radiation dose reduction in a multireader study. Radiology. 2015; 275(3):735–745. [PubMed: 25751228]

11. Solomon J, Wilson J, Samei E. Characteristic image quality of a third generation dual-source MDCT scanner: noise, resolution, and detectability. Med Phys. 2015; 42(8):4941–4953. [PubMed: 26233220]

12. Yu L, Leng S, Chen L, Kofler JM, Carter RE, McCollough CH. Prediction of human observer performance in a 2-alternative forced choice low-contrast detection task using channelized Hotelling observer: impact of radiation dose and reconstruction algorithms. Med Phys. 2013; 40(4):041908. [PubMed: 23556902]

13. Minehiro K, Takata T, Hayashi H, et al. Phantom study on dose reduction using iterative reconstruction in low-dose computed tomography for lung cancer screening [in Japanese]. Nihon Hoshasen Gijutsu Gakkai Zasshi. 2015; 71(12):1201–1208. [PubMed: 26685831]

14. von Falck C, Bratanova V, Rodt T, et al. Influence of sinogram affirmed iterative reconstruction of CT data on image noise characteristics and low-contrast detectability: an objective approach. PLoS One. 2013; 8(2):e56875. [PubMed: 23468886]

15. Holmquist F, Nyman U, Siemund R, Geijer M, Söderberg M. Impact of iterative reconstructions on image noise and low-contrast object detection in low kVp simulated abdominal CT: a phantom study. Acta Radiol. 2016; 57(9):1079–1088. [PubMed: 26663036]

16. Xu J, Fuld MK, Fung GS, Tsui BM. Task-based image quality evaluation of iterative reconstruction methods for low dose CT using computer simulations. Phys Med Biol. 2015; 60(7):2881–2901. [PubMed: 25776521]

17. Ellmann S, Kammerer F, Brand M, et al. A novel pairwise comparison-based method to determine radiation dose reduction potentials of iterative reconstruction algorithms, exemplified through circle of Willis computed tomography angiography. Invest Radiol. 2016; 51(5):331–339. [PubMed: 26741892]

18. Baumueller S, Hilty R, Nguyen TD, Weder W, Alkadhi H, Frauenfelder T. Influence of sinogram-affirmed iterative reconstruction on computed tomography-based lung volumetry and quantification of pulmonary emphysema. J Comput Assist Tomogr. 2016; 40(1):96–101. [PubMed: 26466107]

19. Fletcher JG, Hara AK, Fidler JL, et al. Observer performance for adaptive, image-based denoising and filtered back projection compared to scanner-based iterative reconstruction for lower dose CT enterography. Abdom Imaging. 2015; 40(5):1050–1059. [PubMed: 25725794]

20. Fletcher JG, Yu L, Li Z, et al. Observer performance in the detection and classification of malignant hepatic nodules and masses with CT image-space denoising and iterative reconstruction. Radiology. 2015; 276(2):465–478. [PubMed: 26020436]
21. Gandhi NS, Baker ME, Goenka AH, et al. Diagnostic accuracy of CT enterography for active inflammatory terminal ileal Crohn disease: comparison of full-dose and half-dose images reconstructed with FBP and half-dose images with SAFIRE. Radiology. 2016; 280(2):436–445. [PubMed: 27077382]
22. Pontana F, Billard AS, Duhamel A, et al. Effect of iterative reconstruction on the detection of systemic sclerosis-related interstitial lung disease: clinical experience in 55 patients. Radiology. 2016; 279(1):297–305. [PubMed: 26583761]
23. Remer EM, Herts BR, Primak A, et al. Detection of urolithiasis: comparison of 100% tube exposure images reconstructed with filtered back projection and 50% tube exposure images reconstructed with sinogram-affirmed iterative reconstruction. Radiology. 2014; 272(3):749–756. [PubMed: 24814177]
24. Takx RA, Schoepf UJ, Moscariello A, et al. Coronary CT angiography: comparison of a novel iterative reconstruction with filtered back projection for reconstruction of low-dose CT—initial experience. Eur J Radiol. 2013; 82(2):275–280. [PubMed: 23177187]
25. Winklehner A, Karlo C, Puippe G, et al. Raw data–based iterative reconstruction in body CTA: evaluation of radiation dose saving potential. Eur Radiol. 2011; 21(12):2521–2526. [PubMed: 21822785]
26. Yu L, Fletcher JG, Shiung M, et al. Radiation dose reduction in pediatric body CT using iterative reconstruction and a novel image-based denoising method. AJR Am J Roentgenol. 2015; 205(5): 1026–1037. [PubMed: 26496550]
27. International Commission on Radiation Units and Measurements. ICRU report 54: medical imaging—the assessment of image quality. Bethesda, Md: International Commission on Radiation Units and Measurements; 1995.
28. Hernandez-Giron I, Geleijns J, Calzado A, Veldkamp WJ. Automated assessment of low contrast sensitivity for CT systems using a model observer. Med Phys. 2011; 38(Suppl 1):S25. [PubMed: 21978115]
29. Leng S, Chen B, Vrieze T, et al. Construction of realistic phantoms from patient images and a commercial three-dimensional printer. J Med Imaging (Bellingham). 2016; 3(3):033501. [PubMed: 27429998]
30. Solomon, J., Ba, A., Diao, A., et al. Design, fabrication, and implementation of voxel-based 3D printed textured phantoms for task-based image quality assessment in CT. In: Kontos, D.Flohr, TG., Lo, JY., editors. Proceedings of SPIE: medical imaging 2016—physics of medical imaging. Vol. 9783. Bellingham, Wash: International Society for Optics and Photonics; 2016. p. 978328
31. Solomon, J., Samei, E. Are uniform phantoms sufficient to characterize the performance of iterative reconstruction in CT?. In: Nishikawa, RM., Whiting, BR., editors. Proceedings of SPIE: medical imaging 2013—title. Vol. 8668. Bellingham, Wash: International Society for Optics and Photonics; 2013. p. 86684M
32. Solomon J, Samei E. Quantum noise properties of CT images with anatomical textured backgrounds across reconstruction algorithms: FBP and SAFIRE. Med Phys. 2014; 41(9):091908. [PubMed: 25186395]
33. American Association of Physicists in Medicine. AAPM report no. 204. College Park, Md: American Association of Physicists in Medicine; 2011. Size-specific dose estimates (SSDE) in pediatric and adult body CT examinations.
34. Kachelriess M, Watzke O, Kalender WA. Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT. Med Phys. 2001; 28(4):475–490. [PubMed: 11339744]
35. Stierstorfer K, Rauscher A, Boese J, Bruder H, Schaller S, Flohr T. Weighted FBP: a simple approximate 3D FBP algorithm for multislice spiral CT with good dose usage for arbitrary pitch. Phys Med Biol. 2004; 49(11):2209–2218. [PubMed: 15248573]
36. Solomon J, Samei E. A generic framework to simulate realistic lung, liver and renal pathologies in CT imaging. Phys Med Biol. 2014; 59(21):6637–6657. [PubMed: 25325156]

37. Shuman WP, Chan KT, Busey JM, et al. Standard and reduced radiation dose liver CT images: adaptive statistical iterative reconstruction versus model-based iterative reconstruction—comparison of findings and image quality. Radiology. 2014; 273(3):793–800. [PubMed: 25170546]

38. Goenka AH, Herts BR, Obuchowski NA, et al. Effect of reduced radiation exposure and iterative reconstruction on detection of low-contrast low-attenuation lesions in an anthropomorphic liver phantom: an 18-reader study. Radiology. 2014; 272(1):154–163. [PubMed: 24620913]

39. Schindera ST, Odedra D, Raza SA, et al. Iterative reconstruction algorithm for CT: can radiation dose be decreased while low-contrast detectability is preserved? Radiology. 2013; 269(2):511–518. [PubMed: 23788715]

40. Robins, M., Solomon, J., Sahbaee, P., Samei, E. Development and comparison of projection and image space 3D nodule insertion techniques. In: Kontos, D.Flohr, TG., Lo, JY., editors. Proceedings of SPIE: medical imaging 2016—physics of medical imaging. Vol. 9783. Bellingham, Wash: International Society for Optics and Photonics; 2016. p. 97835X

41. Chen B, Leng S, Yu L, Yu Z, Ma C, McCollough C. Lesion insertion in the projection domain: methods and initial results. Med Phys. 2015; 42(12):7034–7042. [PubMed: 26632058]

42. Gang GJ, Lee J, Stayman JW, et al. Analysis of Fourier-domain task-based detectability index in tomosynthesis and cone-beam CT in relation to human observer performance. Med Phys. 2011; 38(4):1754–1768. [PubMed: 21626910]

43. De Crop A, Smeets P, Van Hoof T, et al. Correlation of clinical and physical-technical image quality in chest CT: a human cadaver study applied on iterative reconstruction. BMC Med Imaging. 2015; 15:32. [PubMed: 26286596]

44. McCollough CH, Yu L, Kofler JM, et al. Degradation of CT low-contrast spatial resolution due to the use of iterative reconstruction and reduced dose levels. Radiology. 2015; 276(2):499–506. [PubMed: 25811326]

Advances in Knowledge

- Detection accuracy decreased with decreasing radiation dose ($P < .001$); this effect was relatively mild at high radiation doses (2% decrease in detection accuracy between 100% and 75% radiation dose) and more pronounced at lower radiation doses (5% decrease in detection accuracy between 25% and 12.5% radiation dose).
- The sinogram-affirmed iterative reconstruction (SAFIRE) algorithm at a full radiation dose increased detectability of subtle hypoattenuating liver lesions compared with filtered back projection (FBP) at a full radiation dose ($P = .03$); furthermore, the SAFIRE algorithm at 16% reduced radiation dose (95% confidence interval: ±13%) had equal detectability to FBP at a full radiation dose.
- The SAFIRE algorithm reduced image noise by 53% ± 5 but also decreased lesion contrast by 12% ± 4 and resulted in subtle hypoattenuating liver lesions with boundaries that were 13% ± 10 blurrier compared with FBP.

Implication for Patient Care

- SAFIRE could be used as an effective way to reduce radiation dose by 16% ± 13 and maintain detectability of subtle hypoattenuating liver lesions at multidetector CT.

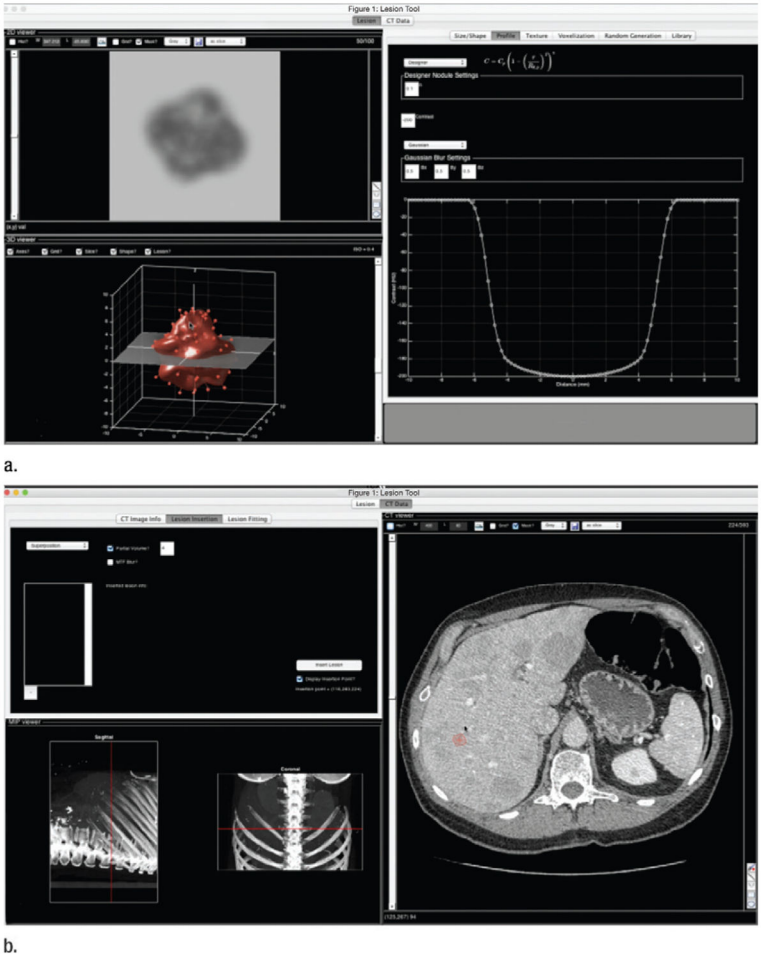

Figure 1.

Graphical user interface of LesionTool software package. **(a)** Tool allows user to create lesion models either by random generation or by fitting to segmented real lesions. **(b)** Tool can also be used to insert lesion models into CT data or to export lesion models for projection-based lesion insertion, as was done for this study.

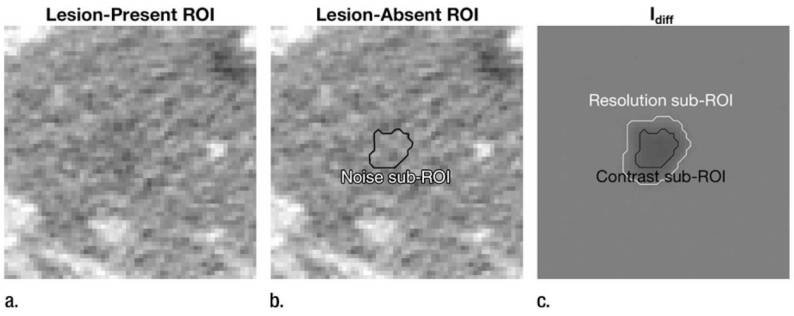

Figure 2.

(a, b) Examples of lesion-present (**a**) and lesion-absent (**b**) ROI at same location. **(c)** Difference of these images, I_{diff}, shows how lesion itself is rendered by reconstruction algorithm. For each lesion in study, noise, contrast, and CNR were measured by using sub-ROIs (within lesion location), as shown. The resolution sub-ROI consisted of the band between the black and white lines, representing the boundary region of the lesion.

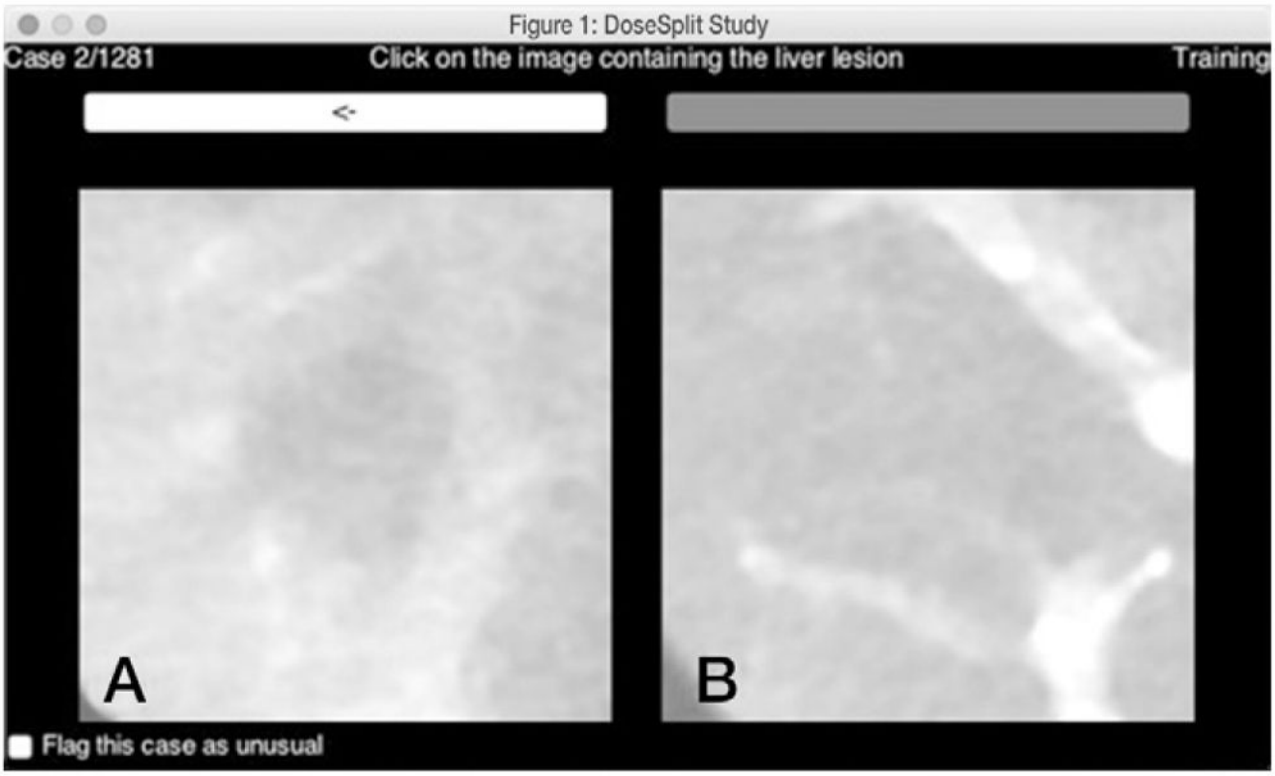

Figure 3.

Graphical user interface of 2AFC experiment shows, *A*, image with lesion and, *B*, image without lesion. For this experiment, the reader was asked to choose which image is most likely to contain the lesion. Reader outcomes (ie, detection or nondetection) were recorded for each trial and the ensemble of outcomes across all combinations of radiation dose level ($n = 6$), reconstruction algorithm ($n = 2$), reader ($n = 16$), and lesion ($n = 105$) were input into a generalized linear mixed effects statistical model to examine how these factors affected detection accuracy.

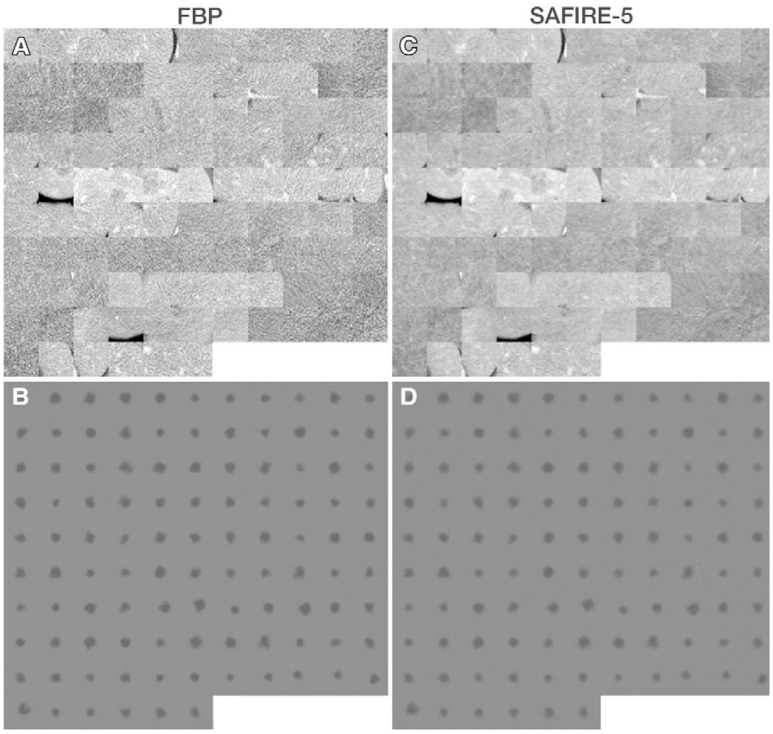

Figure 4.

Montages of ROI images used as basis of comparing FBP with SAFIRE. *A, C,* Montages of CT images show lesions within natural background. *B, D,* Montages show lesions themselves with background subtracted, which represents how lesion appears after being transferred through reconstruction algorithm. Images were obtained from 12.5% radiation dose level scans.

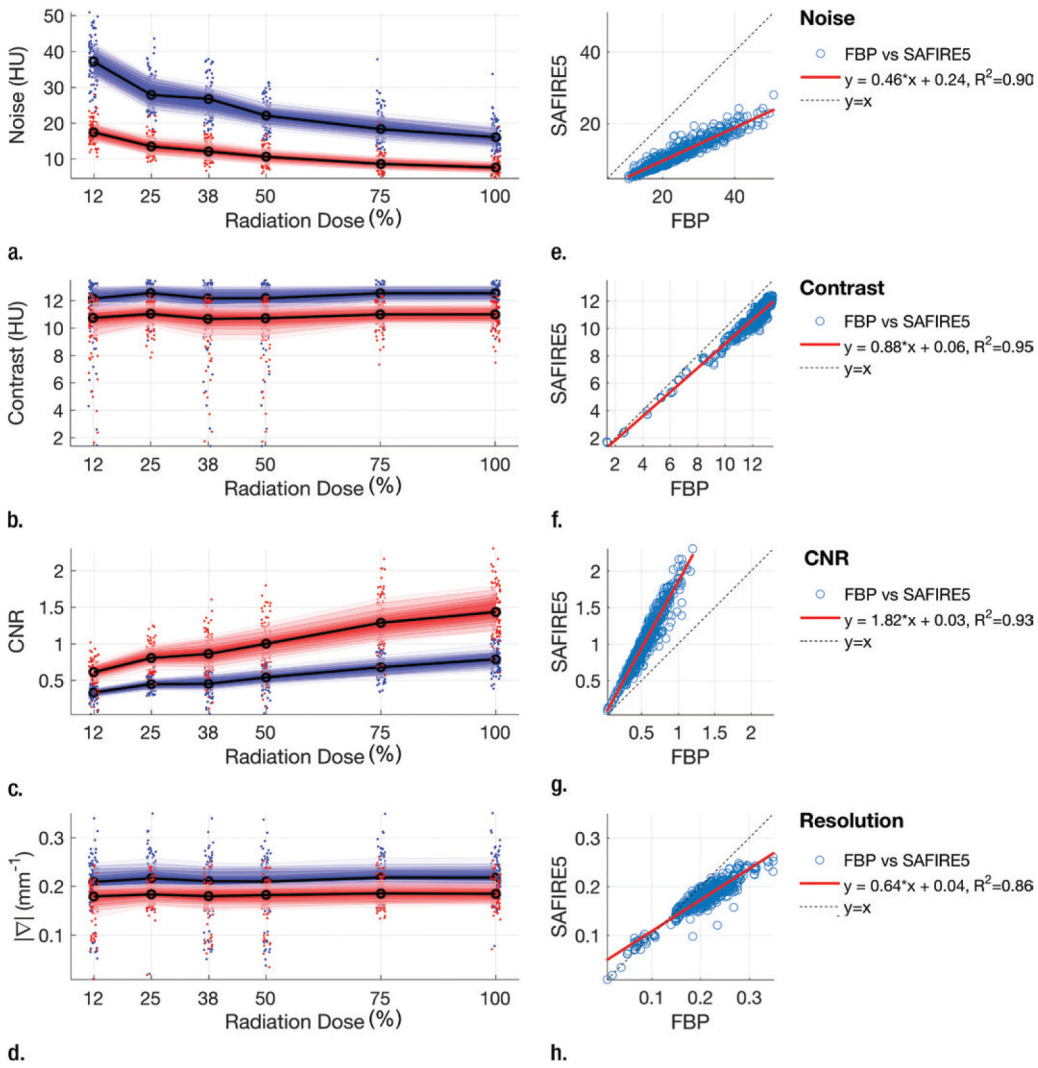

Figure 5.

(a–d) Fan chart plots show noise (**a**), contrast (**b**), CNR (**c**), and average relative gradient magnitude (**d**) as a function of radiation dose. Dots represent individual lesions (blue = FBP, red = SAFIRE), shaded regions represent percentiles of data from 15% (darkest) to 85% (lightest), and solid lines represent medians. **(e–h)** Corresponding linear regression plots show how those factors compare between FBP and SAFIRE images. Data points (blue circles) are shown with their regression lines (red line) and diagonal (y=x, dotted line).

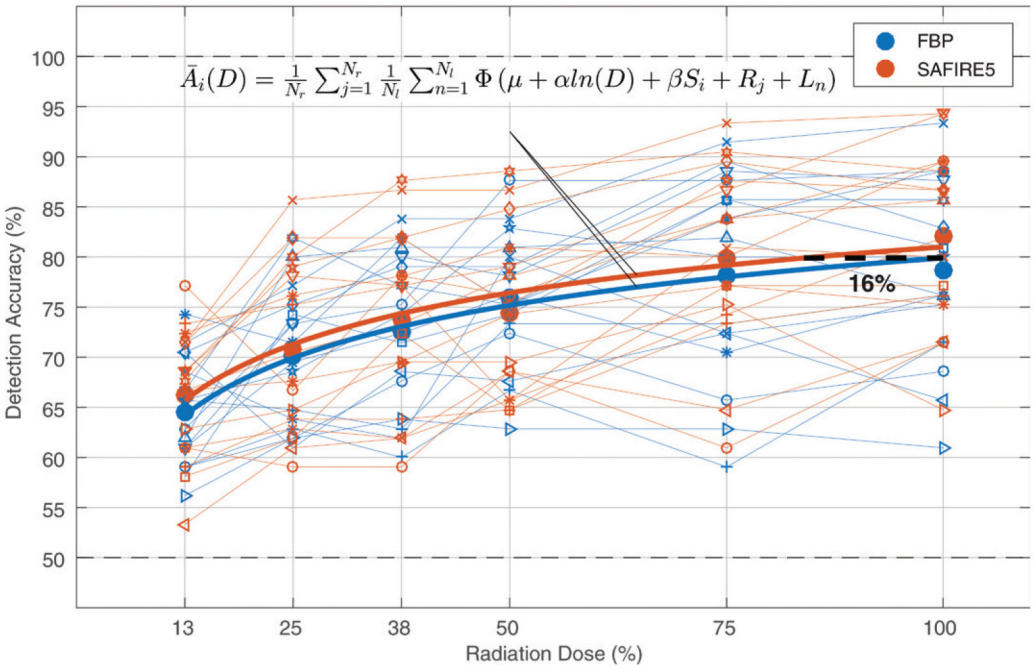

Figure 6.

Results of 2AFC detection experiment show detection accuracy as a function of radiation dose for FBP and SAFIRE. Small markers represent individual readers, large circles are average across readers, and thick solid lines represent detection accuracy averaged across readers and lesions as function of dose, $\bar{A}_i(D)$, as predicted by statistical model. Equation for $\bar{A}_i(D)$ is shown. Dashed line represents radiation dose reduction potential of SAFIRE.

Table 1

Scan Settings for the Abdominal Protocol

Parameter	Data
Scanner model	Somatom Flash
Scan mode	Helical, dual source
Detector configuration (mm)	128 × 0.6*
Beam width (mm)	38.4
Pitch	0.8
Rotation time (sec)	0.33
Tube current modulation	CareDose4D (average setting)
Quality reference effective milliampere second	Scan 1: tube A =150, tube B = 50; scan 2: tube A = 75, tube B = 25
Tube potential (kV)	120
Section thickness (mm)	5
Reconstruction algorithm	FBP, SAFIRE-5
Reconstruction kernel	B31f, I31f

*Note that the flying focal spot was used to achieve a 128-section configuration. The scanner has 64 physical detector rows.

Table 2

Output of the Generalized Mixed Effects Linear Regression Model

A: Fixed effects								
Effect	**Estimate**	**Standard Error**	***t* Statistic**	***P* Value**	**95% CI**			
Intercept (μ)	−0.2238	0.0821	−2.7257	.0064	−0.385, −0.063			
Log (dose) (α)	0.2422	0.0141	17.1352	2.4e-65	0.215, 0.270			
SAFIRE (β)	0.0424	0.0197	2.1559	.0311	0.004, 0.081			
B: random effects								
Effect	N Level	Type	STD					
Reader (R_j)	16	Intercept	0.219					
Lesion (L_n)	105	Intercept	0.309					
c: Fit statistics								
Parameter	*df*	AIC	BIC	Log Likelihood	Deviance	Dev diff	Chi *df*	*P* Value
Constant	3	22068	22091	−11031	22062	…	…	…
Model	5	21771	21811	−10881	21761	300	2	0

Note.—The fitted model shows a significant improvement compared with a constant model according to a difference in deviance test. Note the constant model here had only random terms for the readers and lesions. The deviance shown is −2 × log likelihood. AIC = Akaike information criterion, BIC = Bayesian information criterion, Chi *df* = *df* determined with χ^2 test, Dev diff = deviance difference, STD = standard deviation.